3C

УДК 004.93

С.А. Субботин
Запорожский национальный технический университет
Украина, 69063, г. Запорожье, ул. Жуковского, 64

Критерии сравнения распознающих моделей на основе нейронных сетей и анализ их взаимосвязей

S.A. Subbotin
Zaporizhzhya National Technical University
Ukraine, 69063, Zaporizhzhya, Zhukovskiy st., 64

Criteria for Comparison of Recognition Models Based on Neural Networks and Analysis of Their Mutual Relations

С.А. Субботін
Запорізький національний технічний університет
Україна, 69063, м. Запорожжя, вул. Жуковського, 64

Критерії порівняння розпізнавальних моделей на основі нейронних мереж і аналіз їх взаємозв'язків

Решена актуальная проблема автоматизации построения и анализа распознающих и диагностических моделей на основе нейронных сетей. Предложен комплекс критериев сравнения нейромоделей и методов их синтеза, содержащий в дополнение к традиционно используемым критериям времени и точности критерии, характеризующие логическую прозрачность. Проведены вычислительные эксперименты по исследованию предложенных критериев сравнения нейронных сетей путём решения практических задач.
Ключевые слова: нейронная сеть, распознавание образов, диагностирование, анализ моделей, критерий сравнения.

The actual problem of automation of construction and analysis of recognition and diagnostic models based on neural networks was solved. The set of criteria for comparing of neural network models and synthesis methods was proposed. It contains in addition to the traditionally used time and accuracy criteria the criteria characterizing the logical transparency. Computational experiments on the proposed criteria for comparison of neural networks were conducted by solving practical problems.
Key words: neural network pattern recognition, diagnosis, model analysis, comparison criterion.

Вирішено актуальну проблему автоматизації побудови та аналізу розпізнавальних і діагностичних моделей на основі нейронних мереж. Запропоновано комплекс критеріїв порівняння нейромоделей і методів їх синтезу, що містить на додаток до традиційно використовуваних критеріїв часу і точності критерії, що характеризують логічну прозорість. Проведено обчислювальні експерименти з дослідження запропонованих критеріїв порівняння нейронних мереж шляхом вирішення практичних завдань.
Ключові слова: нейронна мережа, розпізнавання образів, діагностування, аналіз моделей, критерій порівняння.

Введение

В задачах построения распознающих и диагностических моделей широкое распространение на практике получили искусственные нейронные сети (НС), благодаря их способностям обучаться прецедентам [1], [2].

Поскольку на сегодняшний день предложено достаточно много различных типов архитектур НС и еще больше существует вариантов реализации сетей каждого типа,

то большое значение приобретает задача сравнения и выбора такой модели, реализуемой НС, которая удовлетворяла бы целому ряду требований относительно качества функционирования.

Одной из наиболее широко применяемых на практике архитектур НС являются многослойные НС (МНС) прямого распространения сигнала, для которых предложены эффективные методы обучения [1], [3]. Большое видовое разнообразие известных методов обучения МНС также делает актуальным сравнение и отбор наиболее практически полезных методов, что, в свою очередь, требует разработки соответствующих критериев сравнения.

Традиционно используемыми на практике критериями сравнения МНС и методов их построения являются минимум достигаемой ошибки и минимум времени (числа итераций), затрачиваемых на построение модели [1], [3]. Однако при наличии ряда методов, позволяющих получать модели с приемлемыми значениями ошибки и примерно одинаковым временем обучения, выбор наилучшего метода и наилучшей модели оказывается нетривиальной задачей.

Целью данной работы является создание критериев, позволяющих количественно оценивать основные свойства моделей на основе МНС и методов их построения.

Постановка задачи

Пусть мы имеем исходную выборку $X = <x, y>$ – набор S прецедентов о зависимости $y(x)$, $x = \{x^s\}$, $y=\{y^s\}$, $s = 1, 2, ..., S$, характеризующихся набором N входных признаков $\{x_j\}$, $j = 1, 2, ..., N$, где j – номер признака, и выходным признаком y. Каждый s-й прецедент представим как $<x^s, y^s>$, $x^s=\{x^s{}_j\}$, где $x^s{}_j$ – значение j-го входного, а y^s – значение выходного признака для s-го прецедента (экземпляра) выборки, $y^s \in \{1, 2, ..., K\}$, где K – число классов, $K>1$.

МНС можно представить как иерархическую структуру вида:

$$y^s = \psi^{(M,1)}(\varphi^{(M,1)}(w^{(M,i)}, \psi^{(M-1,i)}(...))), \psi^{(0,i)} = x_i^s, w^{(\mu,i)} = \{w_j^{(\mu,i)}\}, i = 1, 2, ..., N_{\mu-1}, \mu = 1, 2, ..., M,$$

где M – число слоёв, N_μ – число узлов (нейронов) в μ-м слое, $\varphi^{(\mu,i)}$ и $\psi^{(\mu,i)}$ – соответственно, дискриминантная (весовая) и активационная функции i-го узла μ-го слоя, $w_j^{(\mu,i)}$ – значение веса j-го входа i-го узла μ-го слоя.

Множество $\{y^{(\mu,i)} = \psi^{(\mu,i)}(\varphi^{(\mu,i)})\}$ определяет структурные блоки МНС, а $\{w_j^{(\mu,i)}\}$ – её параметры.

Тогда задача построения МНС заключается в задании таких её структуры $\{y^{(\mu,i)} = \psi^{(\mu,i)}(\varphi^{(\mu,i)})\}$ и параметров $\{w_j^{(\mu,i)}\}$, которые удовлетворяют комплексу критериев $\Omega = \{\Omega_i\}$, где Ω_i – i-й критерий качества МНС.

Соответственно, для анализа свойств и сравнения методов построения и моделей МНС необходимо определить показатели Ω_i, позволяющие количественно выражать основные свойства МНС и методов их обучения.

Критерии сравнения нейромоделей

Как известно, сходимость итеративных методов обучения МНС, зависит, кроме размерности обучающих данных и начальных значений весов, также от задаваемых максимально допустимой общей ошибки обучения (критерий качества обучения) и количества допустимых циклов обучения (критерий длительности обучения).

Эти два критерия останова целесообразно использовать при сравнении итеративных методов обучения МНС. Зафиксировав поочередно каждый из критериев останова обучения, можно проследить, как для каждого метода обучения МНС при одинаковых обучающих выборках меняются время обучения МНС $t_{об.}$, затраченное количество циклов обучения *epoch* и достигнутая точность (ошибка) E [1], [3].

В [1] для сравнения методов обучения НС предлагается наряду с количеством циклов обучения использовать такой критерий, как N_v – количество дополнительных переменных для организации вычислительного процесса. Под дополнительными здесь понимаются переменные, необходимые для сохранения промежуточных результатов вычислений при программной реализации метода обучения. Предпочтение следует отдавать тем методам, которые при обучении НС требуют малого числа дополнительных переменных, что связано с ограничением ресурсов вычислительных средств.

Достаточно важной характеристикой любой распознающей системы является ее сложность. Исходя из биологической специфики НС, в [2] применительно к ним предлагается использовать следующую терминологию:

– структурная (статическая) сложность системы – описывает построение системы из составляющих ее подсистем;

– вычислительная сложность (сложность управления) – мера вычислительных ресурсов, необходимых для детального расчета особенностей динамики системы, описывает выполняемые операции обработки информации и характеризует практическую сложность детального понимания поведения системы – может быть сведена к зависимости вычислительных возможностей (ресурсов вычислительной системы), необходимых для моделирования поведения системы, от размера задачи.

Структурную сложность нейросетевой модели, основными структурными элементами которой являются нейроны, будем характеризовать количеством нейронов $N_{н.}$, которое для МНС определяется по формуле: $N_{н.} = \sum_{\mu=1}^{M} N_{\mu}$.

Вычислительную сложность нейросетевой модели будем характеризовать числом используемых ячеек памяти $N_{п.м.}$ и количеством времени работы $t_{р.}$, затрачиваемым на расчет значений выходов модели при заданных входах для одного экземпляра, либо количеством времени работы НС для всех экземпляров обучающей выборки $t_{р.о.}$ или тестовой выборки $t_{р.т.}$.

Вычислительную сложность метода синтеза нейросетевой модели будем характеризовать числом используемых ячеек памяти $N_{п.а.}$ и количеством времени, затрачиваемым на обучение модели, $t_{об.}$.

Очевидно, что число используемых ячеек памяти $N_{п.а.}$ будет равно сумме числа ячеек, необходимых для хранения обучающей выборки, числа ячеек, необходимых для хранения переменных модели (в случае НС – весов и порогов), и числа ячеек, необходимых для хранения N_v дополнительных переменных: $N_{п.а.} = \eta_{в.} S\left(N + N_M\right) + \eta_w N_w + \eta_v N_v$, где $\eta_{в.}$ – число ячеек памяти для хранения одного элемента данных обучающей выборки; $S(N+N_M)$ – число элементов данных (размерность) обучающей выборки, включая количество элементов данных целевых признаков; η_w – число ячеек памяти, необходимых для хранения одного веса сети; N_w – количество весов и порогов сети; η_v – число ячеек памяти, необходимых для хранения одной дополнительной переменной метода обучения.

Таких традиционно используемых при сравнении НС и методов их обучения показателей, как точность (ошибка) аппроксимации сетью обучающей или тестовой выборки и время, затрачиваемое на обучение сети и (или) работу обученной сети оказывается недостаточно, поскольку даже несколько НС одного и того же типа, с одинаковой топологией, обученные на одних и тех же данных с помощью одного и того же метода обучения и обеспечивающие почти одинаковую точность аппроксимации и сравнимое время обучения и работы, из-за разных начальных состояний (разных значений весов перед обучением) могут обладать существенно разным качеством, что связано с уровнем логической прозрачности полученной нейросетевой модели, избыточностью памяти и связей сети, качеством аппроксимации обучающих данных.

Поэтому в дополнение к рассмотренным ранее критериям сравнения НС, необходимо разработать комплекс частных и обобщенных критериев, позволяющих количественно оценивать вышеперечисленные свойства нейросетевых моделей, сравнивать их и принимать решение о выборе более оптимальной модели для решения поставленной задачи.

Как было отмечено выше, НС обладают памятью, реализуемой весами. Чем меньше память сети, тем меньше образов она может запомнить, но если две сети с разным объемом памяти обеспечивают требуемую точность распознавания (оценивания), то сеть с меньшей памятью, очевидно, будет проявлять лучшие обобщающие свойства.

Избыточность памяти сети будем характеризовать с помощью коэффициента избыточности для хранения обучающей выборки:

$$K_I = \frac{N_w}{SN},\ N_w > 1,\ S > 0,\ N > 0.$$

Если $K_I > 1$, то память сети избыточна (размерность памяти сети больше размерности обучающей выборки). Если $K_I = 1$, то сеть может запомнить всю обучающую выборку (размерность памяти сети равна размерности обучающей выборки). Если $K_I < 1$, то сеть не сможет в точности запомнить всю обучающую выборку (размерность памяти сети меньше размерности обучающей выборки), однако при этом сеть будет проявлять обобщающие и аппроксимирующие способности.

В случае если количество весов, равных нулю, в НС $N_{w=0}$ велико, то K_I будем определять как:

$$K_I = \frac{N_w - N_{w=0}}{SN}.$$

Поскольку НС являются вычислительными устройствами, весьма важными характеристиками являются сложность их аппаратной или программной реализации, а также время работы в процессе распознавания одного экземпляра данных.

Вычислительную сложность i-го нейрона μ-го слоя $T^{(\mu,i)}$ будем определять как: $T^{(\mu,i)} = N^{(\mu,i)}(T_c^{(\mu,i)} + T_\varphi^{(\mu,i)}) + T_\psi^{(\mu,i)}$, где $N^{(\mu,i)}$ –число входов i-го нейрона μ-го слоя; $T_c^{(\mu,i)}$ – вычислительная сложность одного синапса i-го нейрона μ-го слоя; $T_\varphi^{(\mu,i)}$ – вычислительная сложность дискриминантной функции i-го нейрона μ-го слоя для обработки двух аргументов; $T_\psi^{(\mu,i)}$ – вычислительная сложность функции активации i-го нейрона μ-го слоя. Заметим, что, если $\psi^{(\mu,i)}(x) = x$, то $T_\psi^{(\mu,i)}$ можно положить равной нулю. Вычислительные сложности синапсов, дискриминантных функций и функций активации определяют с учетом условий реализации сети (на разных ЭВМ с разной скоростью работы затраты времени на вычисления будут разными).

Вычислительная сложность сети прямого распространения при последовательной реализации вычислений (при программной реализации на ЭВМ с последовательной организацией вычислений) T_1 определяется как :

$$T_1 = \sum_{\mu=1}^{M} \sum_{i=1}^{N_\mu} T^{(\mu,i)} .$$

Вычислительную сложность сети прямого распространения при параллельной реализации вычислений (например, при аппаратной реализации сети) T_2 будем определять по формуле:

$$T_2 = \sum_{\mu=1}^{M} \max_{i=1,2,\ldots,N_\mu} \left\{T^{(\mu,i)}\right\}.$$

Логическая прозрачность сети в значительной степени зависит от общего количества связей в сети и количества связей, соединяющих конкретные нейроны (количества входов нейронов скрытых слоев). Чем меньше связей, тем проще сеть и тем удобнее она для анализа и интерпретации человеком.

Коэффициент разреженности связей сети прямого распространения определим по формуле:

$$K_R = \frac{N_{w=0}}{\sum_{\mu=1}^{M} N_{(\mu-1)} N_\mu}, \; N_0 = N,$$

где $0 \le N_{w=0} \le \sum_{\mu=1}^{M} N_{(\mu-1)} N_\mu$.

Коэффициент связанности МНС определим как:

$$K_C = 1 - K_R = 1 - \frac{N_{w=0}}{\sum_{\mu=1}^{M} N_{(\mu-1)} N_\mu} .$$

Для коэффициентов связанности и разреженности связей сети прямого распространения характерны следующие свойства: $K_R + K_C = 1$; $0 \le K_R \le 1$;$0 \le K_C \le 1$.

Коэффициент средней связанности нейронной сети прямого распространения K_M показывает среднее количество входов нейронов всех слоев, кроме первого:

$$K_M = \frac{1}{\sum_{\mu=2}^{M} N_\mu} \sum_{\mu=2}^{M} \sum_{i=1}^{N_\mu} \left(N^{(\mu,i)} - N_{w=0}^{(\mu,i)}\right),$$

где $N_{w=0}^{(\mu,i)}$ – число весов i-го нейрона μ-го слоя сети, равных нулю.

Чем больше в сети единичных синапсов (связей, веса которых равны по модулю единице), тем проще ее реализация (прежде всего аппаратная) и удобнее анализ человеком.

Долю единичных синапсов в сети прямого распространения будем рассчитывать по формуле:

$$K_L = \frac{N_{w=1}}{\sum_{\mu=1}^{M} N_{(\mu-1)} N_\mu}, \; N_0 = N,$$

где $N_{w=1}$ – число весов сети, по модулю равных единице: $0 \le N_{w=1} \le \sum_{\mu=1}^{M} N_{(\mu-1)} N_\mu$.

Долю неединичных синапсов в сети прямого распространения K_N определим из выражения:

$$K_N = 1 - K_L = 1 - \frac{N_{w=1}}{\sum_{\mu=1}^{M} N_{(\mu-1)} N_\mu}, \; N_0 = N.$$

Для долей единичных и неединичных синапсов сети прямого распространения характерны следующие свойства: $K_L + K_N = 1$; $0 \le K_L \le 1$; $0 \le K_N \le 1$.

Поскольку логическая прозрачность связей сети зависит в значительной степени от разреженности и простоты связей, будем ее характеризовать коэффициентом, показывающим долю бинарных (нулевых или единичных по модулю) весов в общем количестве весов сети.

Коэффициент логической прозрачности связей сети прямого распространения будем определять из выражения:

$$K_T = K_R + K_L = \frac{N_{w=0} + N_{w=1}}{\sum_{\mu=1}^{M} N_{(\mu-1)} N_\mu}, \; N_0 = N.$$

Коэффициент логической непрозрачности (размытости) связей сети прямого распространения K_S будем рассчитывать как $K_S = 1 - K_T = 1 - \left(K_R + K_L\right)$.

Для коэффициентов логической прозрачности и непрозрачности связей сети прямого распространения характерны следующие свойства: $K_\text{T} + K_S = 1$; $0 \le K_T \le 1$; $0 \le K_S \le 1$.

Для более точной оценки логической прозрачности будем определять логическую прозрачность сети через логическую прозрачность ее элементов.

Коэффициент логической прозрачности i-го нейрона μ-го слоя сети $K_E^{(\mu,i)}$ будет определяться типом используемой функции активации. Для линейной и пороговой функций активации примем: $K_E^{(\mu,i)} = 1$, для всех остальных функций активации $K_E^{(\mu,i)} = 0$.

Коэффициент логической прозрачности МНС K_U будем рассчитываться как:

$$K_U = \frac{\sum_{\mu=1}^{M} \sum_{i=1}^{N_\mu} K_E^{(\mu,i)}}{K_M \sum_{\mu=1}^{M} N_\mu}, K_M \ne 0.$$

Чем больше K_U, тем выше уровень логической прозрачности сети, и, наоборот, чем меньше K_U, тем ниже уровень логической прозрачности сети.

Одной из важнейших характеристик нейросетевых моделей является качество аппроксимации. Качество аппроксимации при одном и том же уровне ошибки тем выше, чем меньше используется весов.

Коэффициент качества аппроксимации нейросетевой модели K_A определим как среднюю долю ошибки, приходящуюся на ненулевые веса сети:

$$K_A = \frac{E}{N_w - N_{w=0}},$$

где E – совокупная ошибка, допускаемая сетью (например, среднеквадратическая ошибка), такая, что $E \le \xi$, где ξ – максимально допустимая ошибка (цель обучения). В качестве ошибки Е можно использовать ошибку по обучающей выборке $E_{\text{об.}}$ или по тестовой выборке $E_{\text{т.}}$.

Вычислительные аспекты реализации расчёта критериев сравнения нейромоделей

Рассмотрим вычислительные аспекты некоторых критериев сравнения методов синтеза нейромоделей.

Поскольку важнейшими характеристиками итеративных методов обучения МНС являются время обучения $t_{об.}$, количество затраченных циклов обучения *epoch* и достигнутая ошибка E, для отбора наилучшего среди градиентных методов обучения МНС целесообразно для различных прикладных задач сравнить методы по каждому из данных критериев, зафиксировав поочередно остальные.

Для неитеративных методов синтеза МНС следует оценивать время обучения $t_{об.}$ и достигнутую точность (ошибку) E.

Значения критериев N_v и $N_{п.а.}$ будут сильно зависеть от особенностей программной реализации каждого метода обучения. При этом для градиентных методов большое значение будет иметь способ вычисления производной целевой функции по весам. В связи с данными обстоятельствами приведенные выше теоретические оценки данных критериев целесообразно на практике заменить на следующие:

$$N_{п.а.} = (M_2 - M_1) + (M_3 - M_2) + (M_4 - M_3) = M_4 - M_1,$$

$$M_4 - M_3 = \eta_v N_v,\ M_2 - M_1 = \eta_{в.} S(N + N_M),\ M_3 - M_2 = \eta_w N_w,$$

где M_1 – размер занятой памяти ЭВМ до загрузки обучающей выборки, M_2 – размер занятой памяти ЭВМ после загрузки обучающей выборки перед созданием переменных модели НС, M_3 – размер занятой памяти ЭВМ после создания переменных модели и перед обучением НС, M_4 – размер занятой памяти ЭВМ перед окончанием процедуры обучения НС. Таким образом, для оценки значений критериев N_v и $N_{п.а.}$ следует измерять размеры занятой памяти ЭВМ M_1, M_3 и M_4.

Критерии $N_{п.а.}$, E, $N_{н.}$, $N_{п.м.}$, $t_{р.о.}$, $t_{р.т.}$, K_I, T_1, T_2, K_R, K_C, K_M, K_L, K_N, K_T, K_S, K_U и K_A необходимо оценивать для моделей, полученных в результате использования как итеративных, так и неитеративных методов. При этом значение критерия $N_{п.м.}$, будем определять следующим образом: $N_{п.м.} = (M_3 - M_2)$.

В том случае, когда в НС доля синапсов с единичными весами велика и зависит только от размерности данных, а не от их значений, а сами единичные веса являются константными и задаются некоторым несложным правилом, эти веса можно не хранить в памяти, используя соответствующее правило в процессе работы модели. Тогда эти веса можно исключить из общего числа весов при определении K_I, получив таким образом формулу для скорректированного коэффициента избыточности K_I':

$$K_I' = \frac{N_w - (N_{w=0} + N_{w=1})}{NS} = K_I - \frac{N_{w=1}}{NS}.$$

В этом случае K_I' будет учитывать только те веса, которые обеспечивают настройку НС на решение конкретной задачи, а фактическая избыточность сети в этом случае окажется существенно меньше.

Эксперименты и результаты

Для исследования практической применимости предложенных критериев, а также выявления их взаимосвязей было разработано программное обеспечение, автоматизирующее расчет комплекса предложенных критериев для нейромоделей. Построение нейромоделей осуществлялось на основе обучающих выборок данных, различной размерности для задач распознавания образов, технического и медицинского диагностирования.

На рис. 1. и рис. 2 приведены экспериментально полученные графики зависимостей между предложенными показателями свойств нейромоделей на комплексе обучающих выборок для различных задач распознавания.

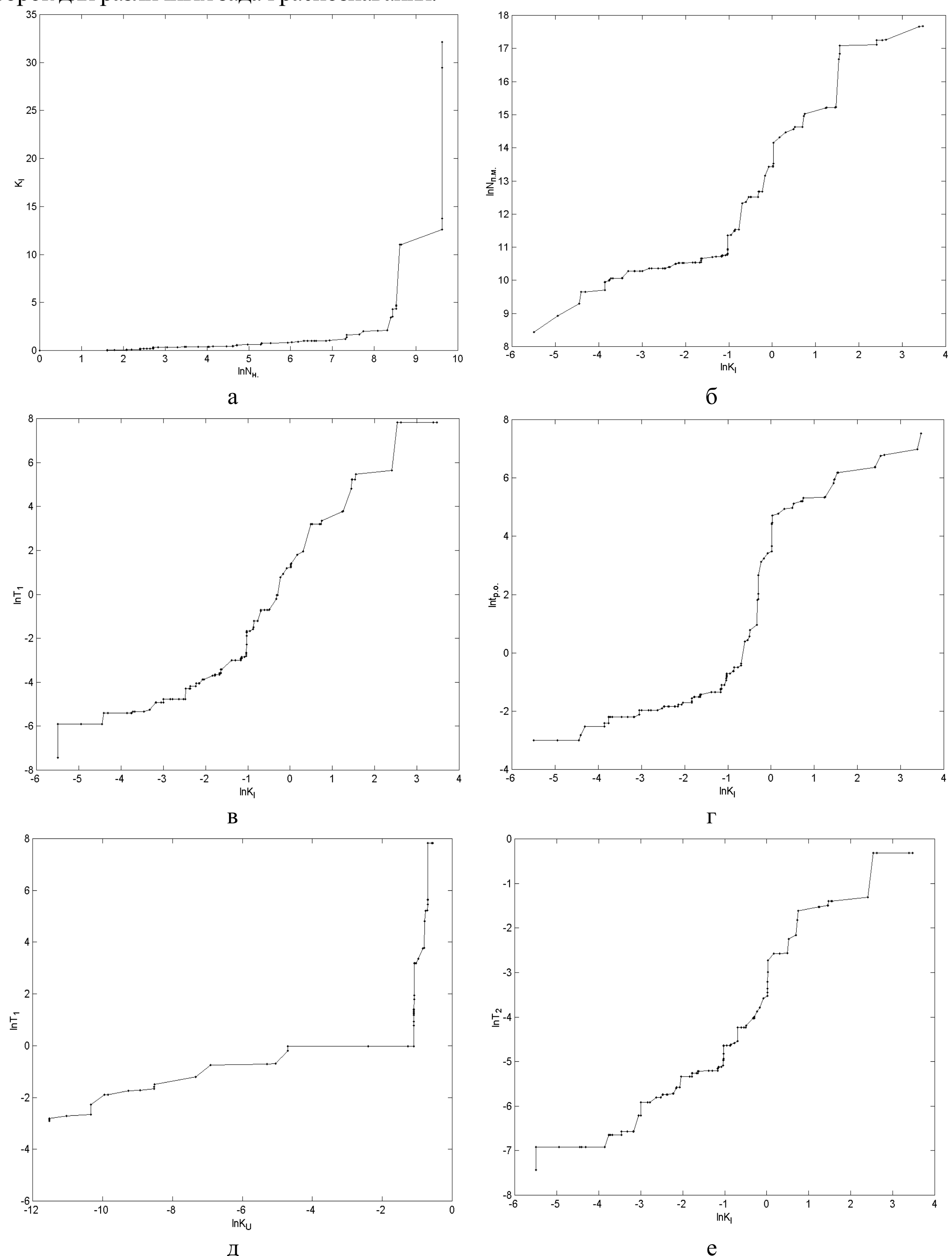


Рисунок 1 – Графики зависимостей: а – K_I от $N_{н.}$ в полулогарифмической системе координат; б – $N_{п.м.}$ от K_I в логарифмической системе координат; в – T_1 от K_I в логарифмической системе координат; г – $t_{р.о.}$ от K_I в логарифмической системе координат; д – T_1 от K_U в логарифмической системе координат; е – T_2 от K_I в логарифмической системе координат

Как видно из рис. 1а, коэффициент избыточности K_I тесно связан с количеством нейронов в сети $N_{н.}$: он возрастает с увеличением и убывает с уменьшением $N_{н.}$.

Объем памяти ЭВМ, занимаемый моделью МНС, $N_{п.м.}$ возрастает с увеличением коэффициента избыточности K_I (рис. 1б). А поскольку ресурсы памяти ЭВМ ограничены, объем занимаемой памяти $N_{п.м.}$ может быть уменьшен либо путем уменьшения $N_{н.}$ и N_w (что, в свою очередь, частично может быть достигнуто за счет уменьшения числа признаков и входов N, количества слоев M и числа нейронов в каждом слое N_μ), либо за счет использования более эффективных структур данных.

На рис. 1в изображён график зависимости времени работы нейромодели при последовательной реализации вычислений от коэффициента избыточности сети. С увеличением коэффициента избыточности НС наблюдается увеличение времени работы нейромодели при последовательной реализации вычислений. Поэтому для сокращения времени T_1 необходимо уменьшить избыточность нейромодели.

На рис. 1г изображён график зависимости времени работы нейромодели для обучающей выборки от коэффициента избыточности нейромодели в логарифмической системе координат. С увеличением избыточности нейромодели наблюдается увеличение времени работы НС для обучающей выборки.

Теоретические оценки времени распознавания экземпляра на ЭВМ для МНС возрастают с увеличением избыточности K_I и логической прозрачности K_U (рис. 1д и рис. 1е).

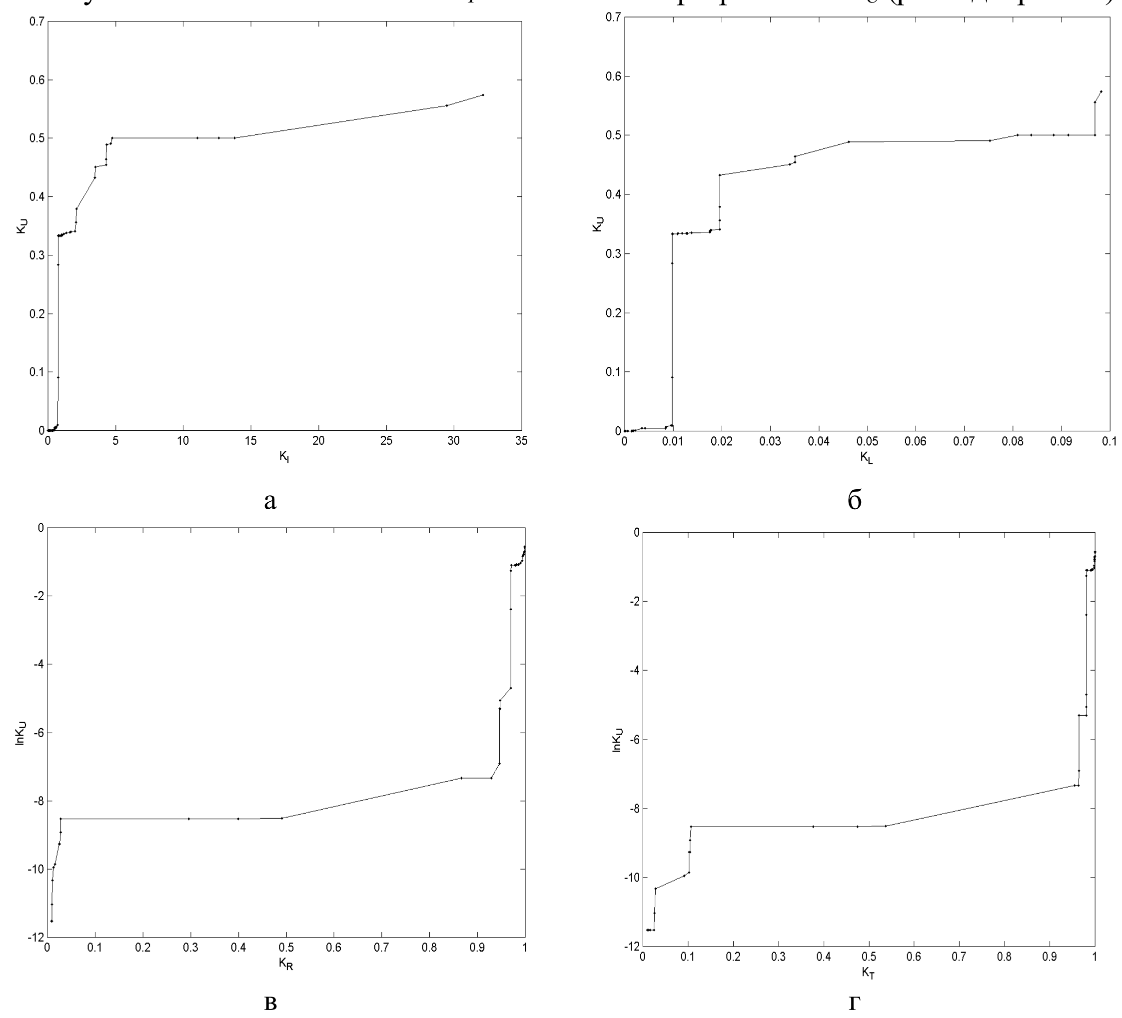


Рисунок 2 – Графики зависимостей K_U от: а – K_I; б – K_L; в – K_R в полулогарифмической системе координат; г – K_T в полулогарифмической системе координат

Как видно из рис. 2, коэффициент логической прозрачности K_U возрастает с увеличением K_I, K_L, K_R и K_T, и наоборот – убывает с их уменьшением. Следовательно, логическая прозрачность МНС будет возрастать с увеличением долей единичных и нулевых весов. При этом существенное увеличение логической прозрачности на практике предполагает повышение избыточности нейромодели, что связано, прежде всего, с обеспечивающейся при этом структуризацией модели.

Проведенные эксперименты подтвердили практическую полезность предложенного комплекса критериев сравнения МНС и позволяют рекомендовать его для использования при решении задач распознавания образов и диагностирования на основе МНС.

Выводы

В статье предложено новое решение актуальной проблемы автоматизации построения и анализа распознающих и диагностических моделей на основе НС посредством разработки критериев сравнения методов построения и моделей МНС.

Научная новизна полученных результатов состоит в том, что предложен комплекс критериев сравнения нейромоделей и методов их синтеза, содержащий в дополнение к традиционно используемым критериям времени и точности критерии, характеризующие логическую прозрачность. Это дает возможность автоматизировать выбор наиболее удобных для последующего использования нейромоделей и методов их синтеза.

Практическая значимость полученных результатов заключается в том, что проведены вычислительные эксперименты по исследованию предложенных критериев сравнения МНС путём решения практических задач распознавания образов и диагностирования, результаты которого подтверждают практическую пригодность предложенного комплекса критериев.

Дальнейшие исследования могут быть сосредоточены на разработке аналогичных предложенным критериям критериев сравнения нейронных сетей с латеральными и обратными связями.

Работа выполнена в рамках госбюджетной научно-исследовательской темы Запорожского национального технического университета «Интеллектуальные информационные технологии автоматизации проектирования, моделирования, управления и диагностирования производственных процессов и систем» (номер гос. регистрации 0112U005350).

Список литературы

1. Круглов В. В. Искусственные нейронные сети: Теория и практика / В. В. Круглов, В. В. Борисов. – М.: Горячая линия-Телеком, 2001. – 382 с.
2. Рамбиди Н. Г. Биомолекулярные нейросетевые устройства / Н. Г. Рамбиди. – М.: ИПРЖР, 2002.– 224 с.
3. Интеллектуальные средства диагностики и прогнозирования надежности авиадвигателей: монография / [В. И. Дубровин, С. А. Субботин, А. В. Богуслаев, В. К. Яценко]. – Запорожье: ОАО "Мотор-Сич", 2003. – 279 с.
4. Интеллектуальные информационные технологии проектирования автоматизированных систем диагностирования и распознавания образов : монография / [С. А. Субботин, Ан. А. Олейник, Е. А. Гофман, С. А. Зайцев, Ал. А. Олейник] ; под ред. С. А. Субботина. – Харьков : ООО «Компания Смит», 2012. – 317 с.
5. Прогрессивные технологии моделирования, оптимизации и интеллектуальной автоматизации этапов жизненного цикла авиационных двигателей : монография / [А. В. Богуслаев, Ал. А. Олейник, Ан. А. Олейник, Д. В. Павленко, С. А. Субботин] ; под ред. Д. В. Павленко, С. А. Субботина. – Запорожье: ОАО "Мотор Сич", 2009. – 468 с.

References

1. Kruglov V. V. Iskusstvennye nejronnye seti: Teorija i praktika / V. V. Kruglov, V. V. Borisov. – M.: Gorjachaja linija-Telekom, 2001. – 382 p.
2. Rambidi N. G. Biomolekuljarnye nejrosetevye ustrojstva / N. G. Rambidi. – M.: IPRZhR, 2002.– 224 c.
3. Intellektual'nye sredstva diagnostiki i prognozirovanija nadezhnosti aviadvigatelej: mono-grafija / [V. I. Dubrovin, S. A. Subbotin, A. V. Boguslaev, V. K. Jacenko]. – Zaporozh'e: OAO "Motor-Sich", 2003. – 279 p.
4. Intellektual'nye informacionnye tehnologii proektirovanija avtomatizirovannyh sistem diagnostirovanija i raspoznavanija obrazov : monografija / [S. A. Subbotin, An. A. Olejnik, E. A. Gofman, S. A. Zajcev, Al. A. Olejnik] ; pod red. S. A. Subbotina. – Har'kov : OOO «Kom-panija Smit», 2012. – 317 p.
5. Progressivnye tehnologii modelirovanija, optimizacii i intellektual'noj avtomatizacii jeta-pov zhiznennogo cikla aviacionnyh dvigatelej : monografija / [A. V. Boguslaev, Al. A. Olejnik, An. A. Olejnik, D. V. Pavlenko, S. A. Subbotin] ; pod red. D. V. Pavlenko, S. A. Subbotina. – Zaporozh'e: OAO "Motor Sich", 2009. – 468 p.

S.A. Subbotin

Criteria for Comparison of Recognition Models Based on Neural Networks and Analysis of Their Mutual Relations

The artificial neural networks have received a widespread usage in practice in the tasks of recognition and diagnostic model building.

The great diversity of known methods of neural network training makes actual the comparison and selection of the most practically useful methods, which in turn requires the development of appropriate comparison criteria .

The aim of this work is to create criteria to evaluate quantitatively the main properties of the models based on multilayer neural networks and methods for their construction.

The paper proposes a new solution of the urgent problems of automation of construction and analysis of recognition and diagnostic models based on neural networks through the development of criteria for comparing methods for constructing and models of multilayer neural networks.

The scientific novelty of the results is that the set of criteria for comparing neural network models and methods for their synthesis is proposed, including, in addition to the traditionally used time and accuracy criteria the criteria characterizing the logical transparency. This makes possible to automate the choice of the most suitable neural network models and methods for their synthesis for the subsequent usage.

To investigate the practical applicability of the proposed criteria, as well as to identify their relationships the software automatising evaluation of the proposed criteria complex for neural models were developed. The neural network model synthesis was based on the training data samples with different dimensionality for pattern recognition technical and medical diagnosis problems.

The practical significance of obtained results is that the computational experiments were conducted to study the proposed criteria for comparison of multilayer neural networks by solving practical problems of pattern recognition and diagnosis, which results confirms the practical relevance of the proposed set of criteria.

The further research may focus on the development of similar to the proposed criteria the comparison criteria of neural networks with lateral and feedback links.

Статья поступила в редакцию 27.12.2013.